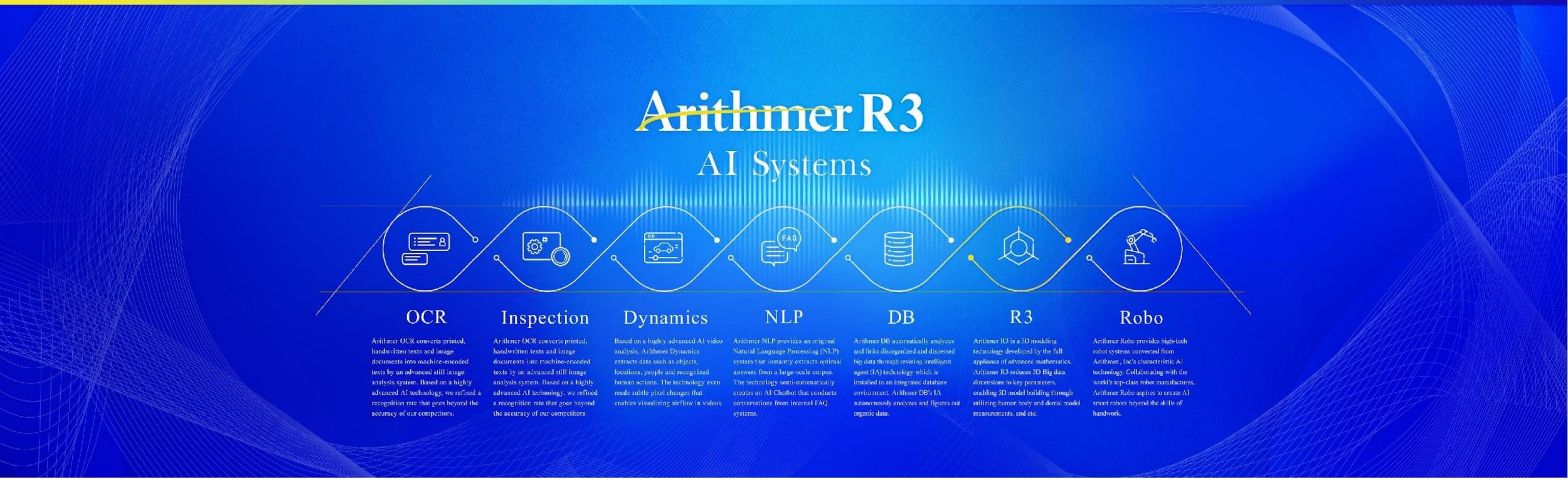


Weakly Supervised Semantic Segmentation in 3D Graph-Structured Point Clouds of Wild Scenes

Arithmer R3 team
Daisuke SATO

https://arxiv.org/abs/2004.12498v2

2021/7/18

©2020 Arithmer.Inc. 当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。 Arithmer Confidential

Point clouds

- 3D sensors have been developed rapidly these days.
  - iPhone/iPad has LiDAR
  - realsense/azure kinect is amazing considering its low price.

- Raw data collected is point cloud
  - Collection of 3D points
  - (x_0, y_0, z_0)
    (x_1, y_1, z_1)
    …
    (x_n, y_n, z_n)

- **Processing point cloud is quite important in robotics/measurement (測量).**

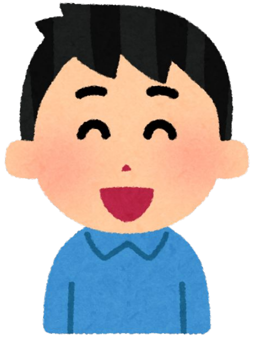
3D Point Cloud

©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。 Arithmer Confidential

Task: semantic segmentation of point clouds

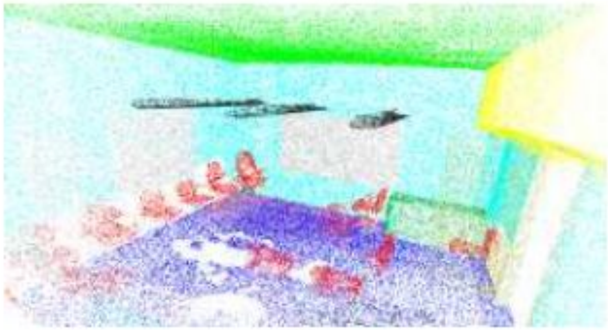
- Classifying every point of 3D point clouds
  - If we perform this task with supervised learning, we need to annotate each point.
  - **Labeling point clouds is super exhausting.**

  - Labeling 2D image is easier.
  - **Let us do weakly supervised learning with 2D projected image.**



3D Point Cloud

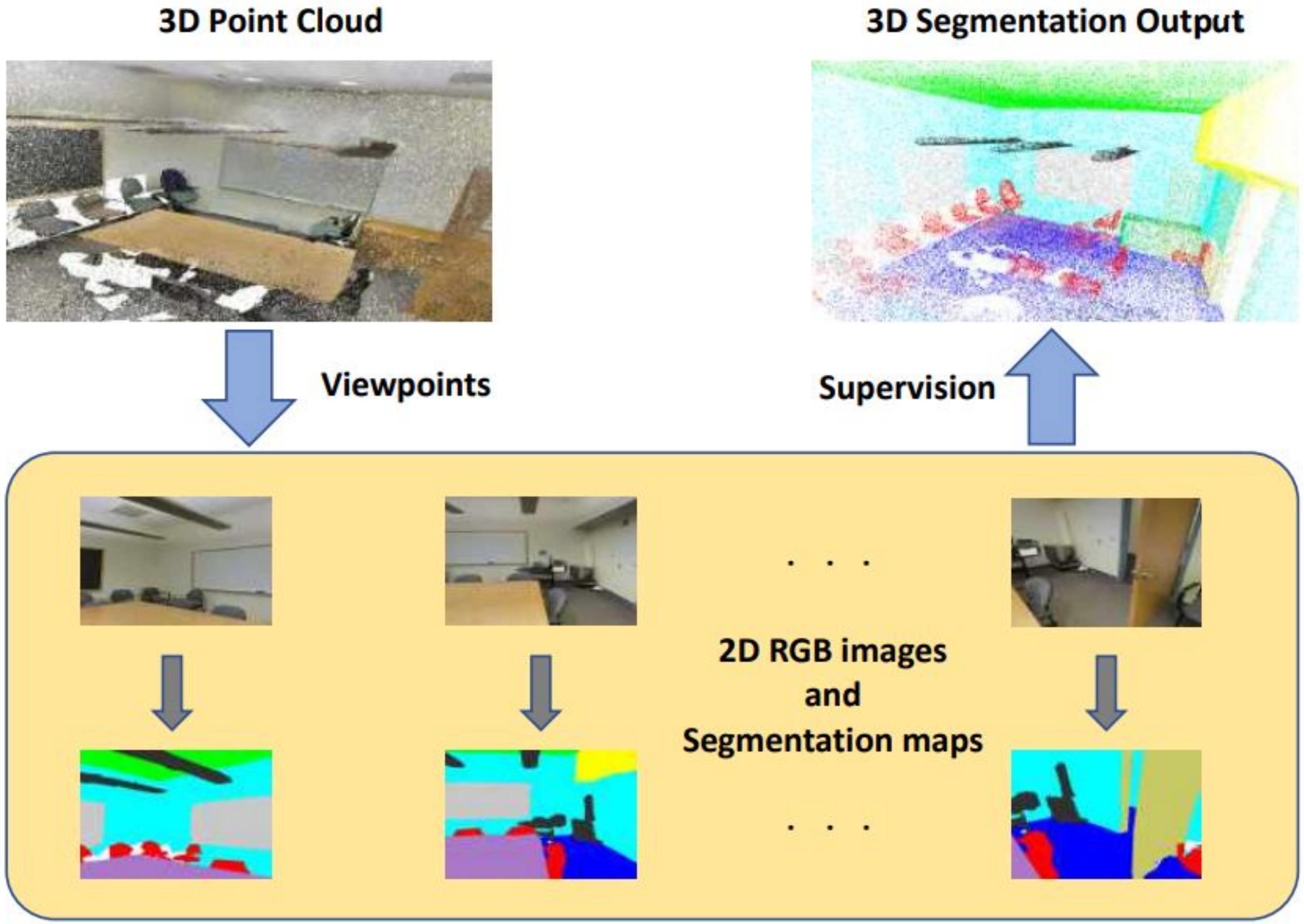
3D Segmentation Output



©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。
Arithmer Confidential

Idea: weakly supervised learning with 2D projected image

Input is
3D point cloud



Output at
inference stage
is 3D segmented
point cloud

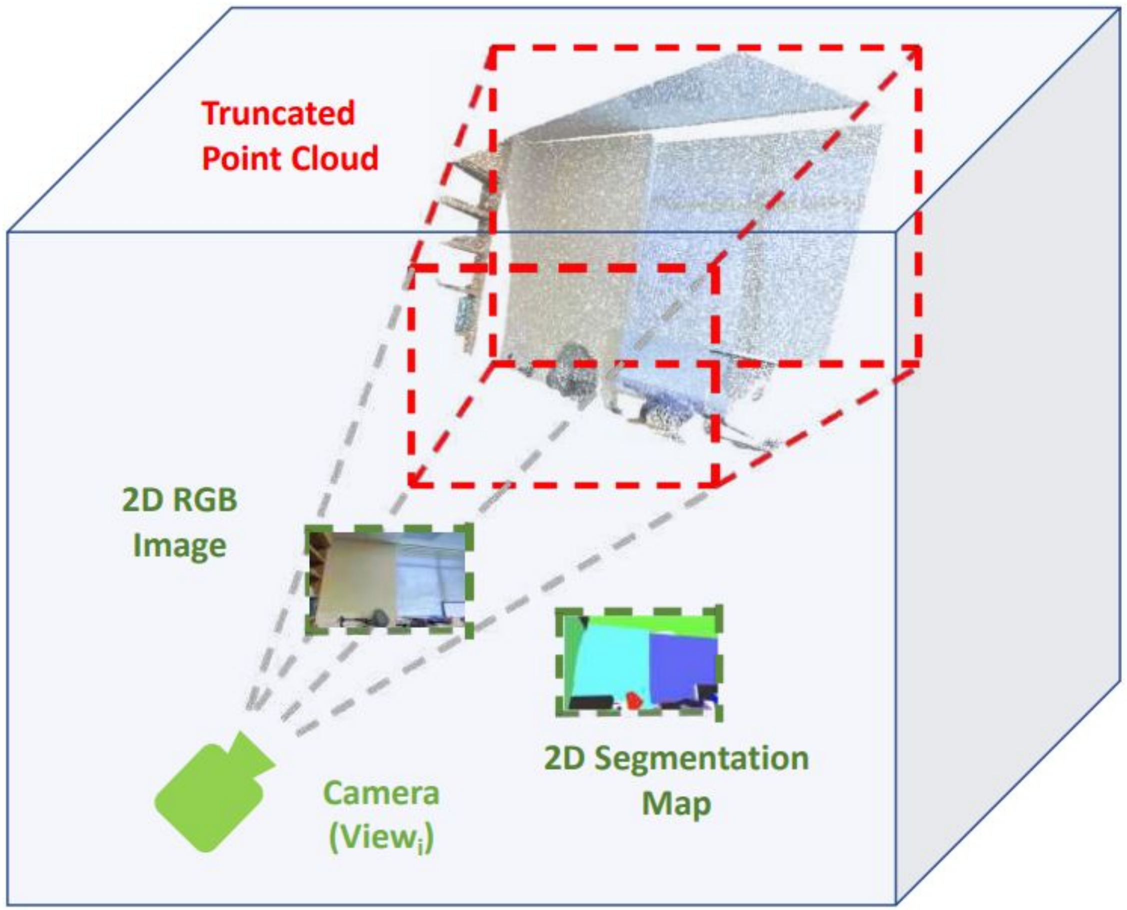
Label at training
stage is
segmented 2D
projected images

©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。
Arithmer Confidential

Note: Projection of 3D pointcloud onto 2D image

- Projection is described by the two parameters:
  - Internal camera parameters
    - Focal length (f)
    - principal point (c)

  - External camera parameters
    - Translation (t)
    - Rotation (R)



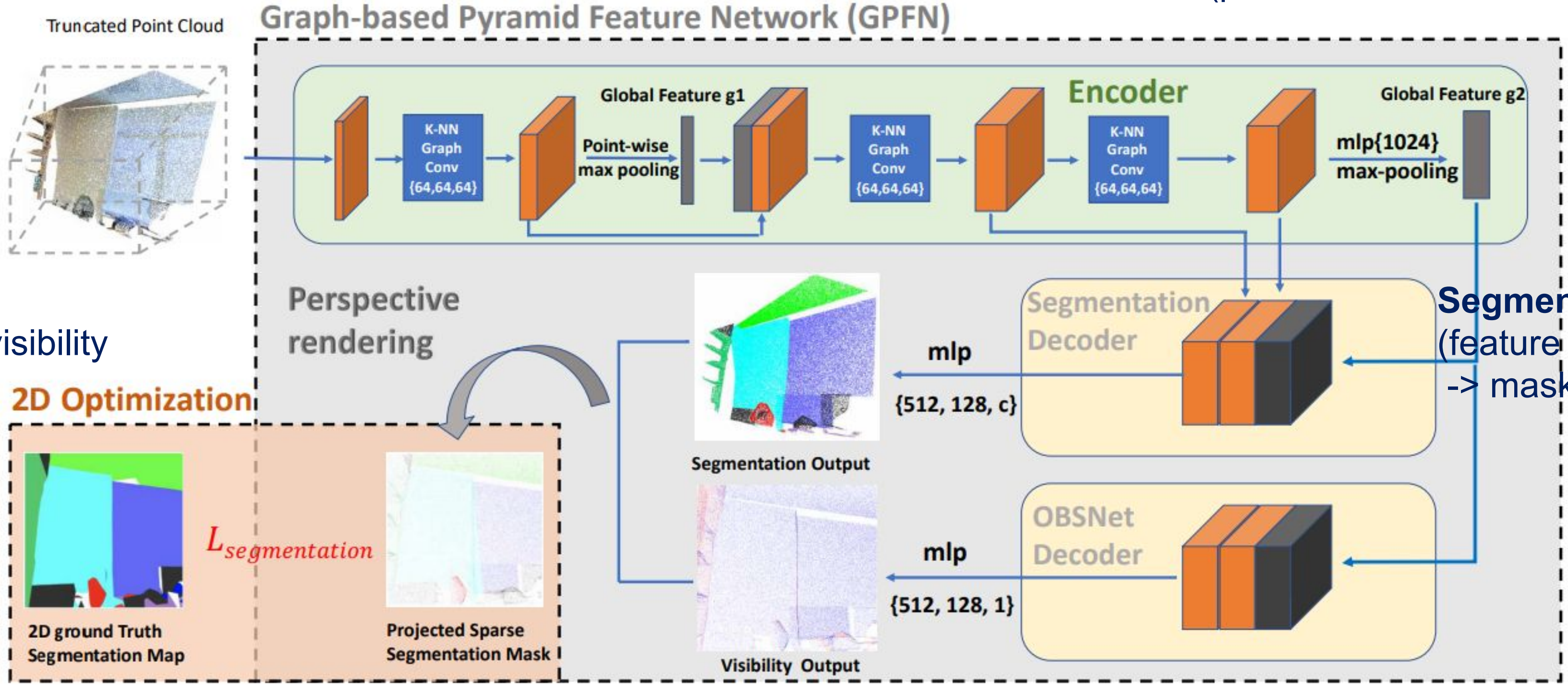
©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。
Arithmer Confidential

Model architecture

Graph convolution encoder
(point cloud -> feature vector)

Input is 3D point cloud

2D projection
(segmented pcd/mask for visibility
-> 2D segmented image)



Segmentation branch
(feature vector
-> mask for segmentation)

$$L_{seg} = -\frac{1}{N}\sum_{i=1}^{N}[p_i \log \hat{p}_i + (1-p_i)\log(1-\hat{p}_i))]$$

Visibility branch
(feature vector -> mask for visibility)

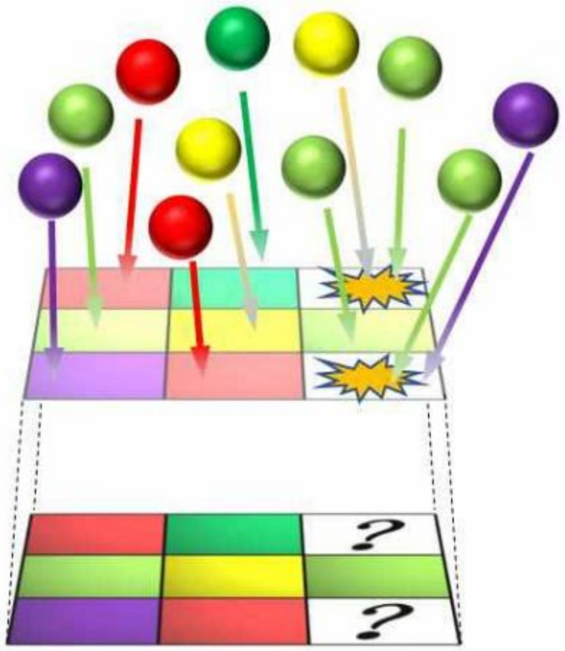
$$L_{vis} = -\frac{1}{M}\sum_{i=1}^{M}[U_i \log \hat{U}_i + (1-U_i)\log(1-\hat{U}_i))]$$

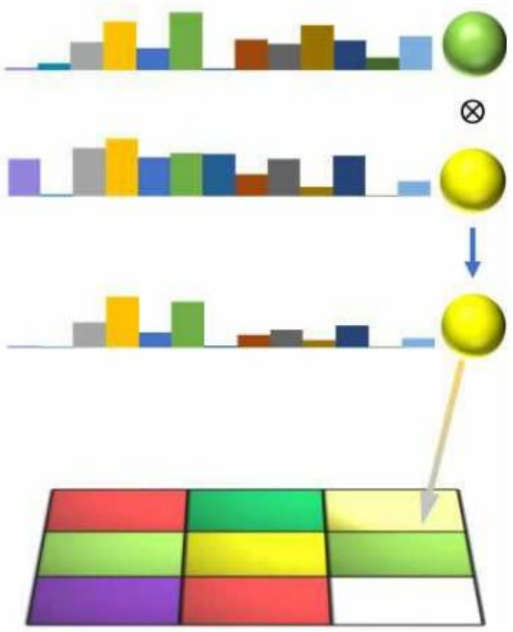
Total loss $L = L_{seg} + \lambda L_{vis}$

©2020 Arithmer.Inc. 当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。 Arithmer Confidential

Perspective rendering

- Problem: Multiple pointclouds can be projected onto a single pixel. Which class should we give the pixel?
- Solution: “Semantic fusion” (sophiscated voting system)





$$p(C_i|x_{grid}) = \prod_{n=1}^{N} p(C_i|x_n),$$

$$p(C_i|x_{grid})_{norm} = p(C_i|x_{grid}) / \sum_{i=1}^{n_{classes}} \prod_{n=1}^{N} p(C_i|x_n),$$

$$p(x_{grid}) = max\{p(C_1|x_{grid}), ..., p(x_{C_{n_{classes}}}|grid)\}.$$

©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。　Arithmer Confidential

Datasets

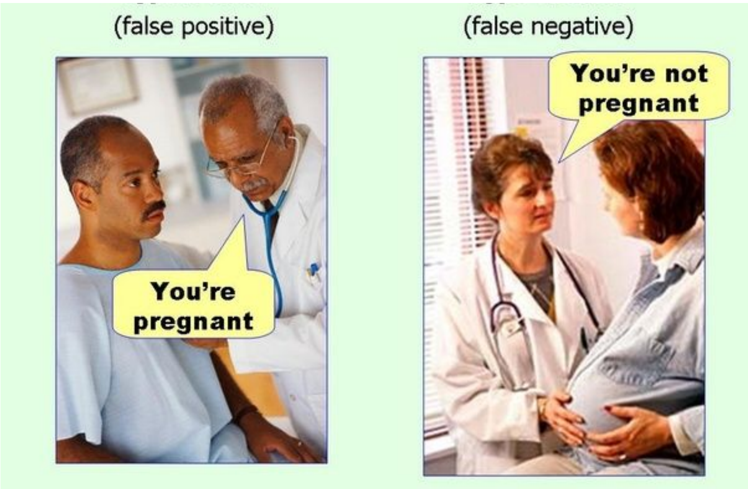
- SUNCG **Synthetic** dataset
  - class number: 40 (furniture)
  - rooms: 404058
    - create 55000 2D rendering sets
- S3DIS **Real-world** dataset
  - class number: 13 (furniture)
  - rooms: 272
    - thousands of viewpoints are provided
- Each point has
  - position: (x, y, z)
  - color: (r, g, b)
- Also normal is computed (u, v, w)

©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。 Arithmer Confidential

Metrics

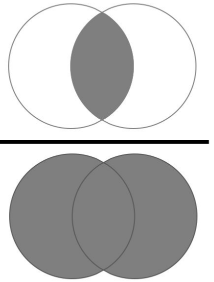
- mean accuracy of total classes ($mAcc$)
- overall accuracy ($oAcc$)

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$



- mean per-class intersection-over-union ($mIoU$)
  - IOU = TP / (TP + FP + FN)

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$



©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。
Arithmer Confidential

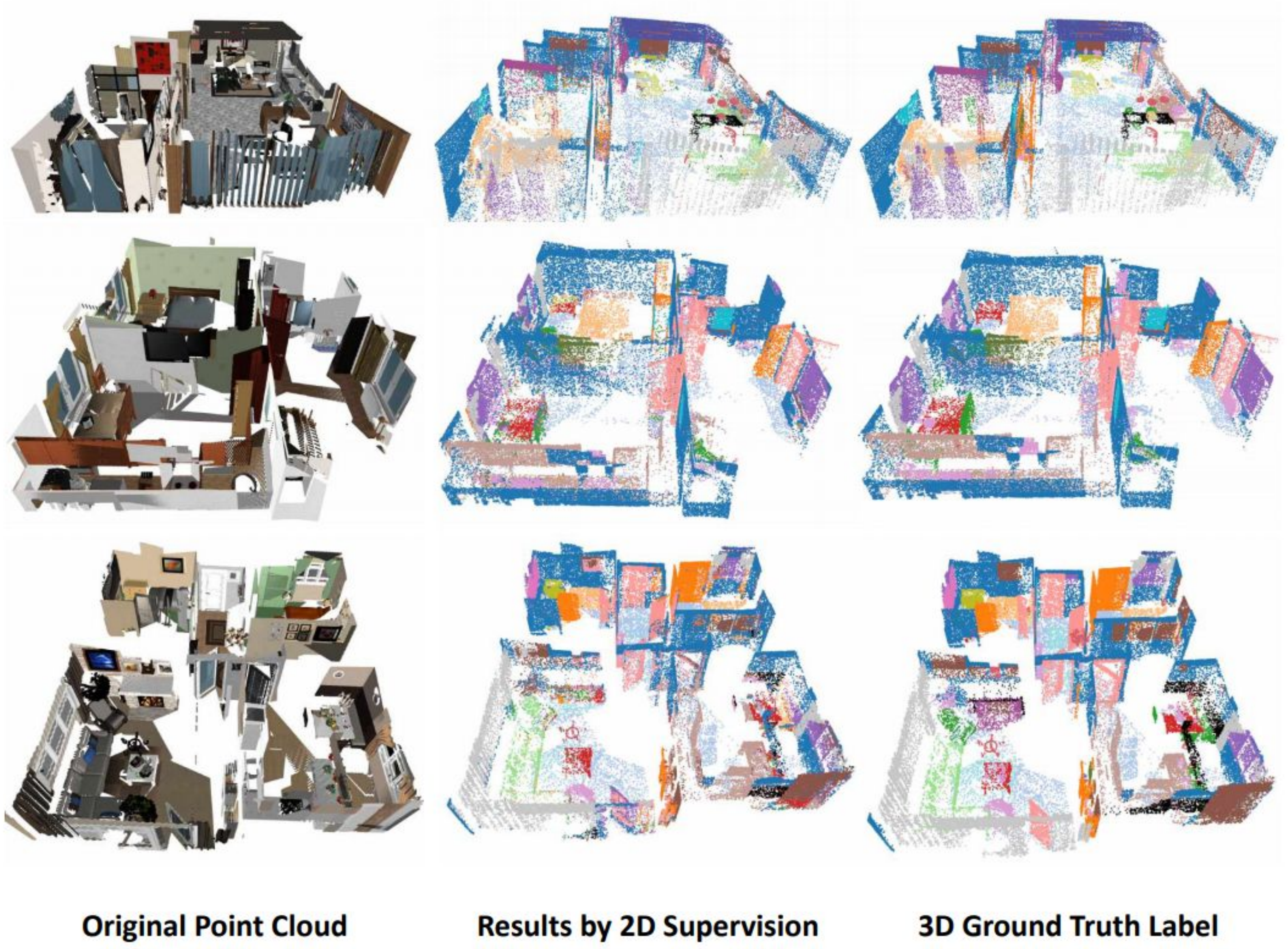
Comparison with other fully-supervised methods

	Method	mAcc(%)	mIoU(%)	oAcc(%)
3D Supervision	*PointNet* [28]	66.2	47.6	78.5
	Engelmann et al. [10]	66.4	49.7	81.1
	PointNet++ [29]	67.1	54.5	81.0
	DGCNN [40]	-	56.1	84.1
	Engelmann et al. [11]	67.8	58.3	84.0
	SPG [19]	73.0	62.1	85.5
2D Supervision	*GPFN with DP (Ours)*	39.2	30.4	53.7
	GPFN with DP w/ D_v (Ours)	59.4	42.7	70.0
	GPFN with PR w/o D_v (Ours)	54.2	39.0	66.8
	GPFN with PR w/ D_v (Ours)	66.5	50.8	79.1

Not so bad even compared with fully-supervised learning.

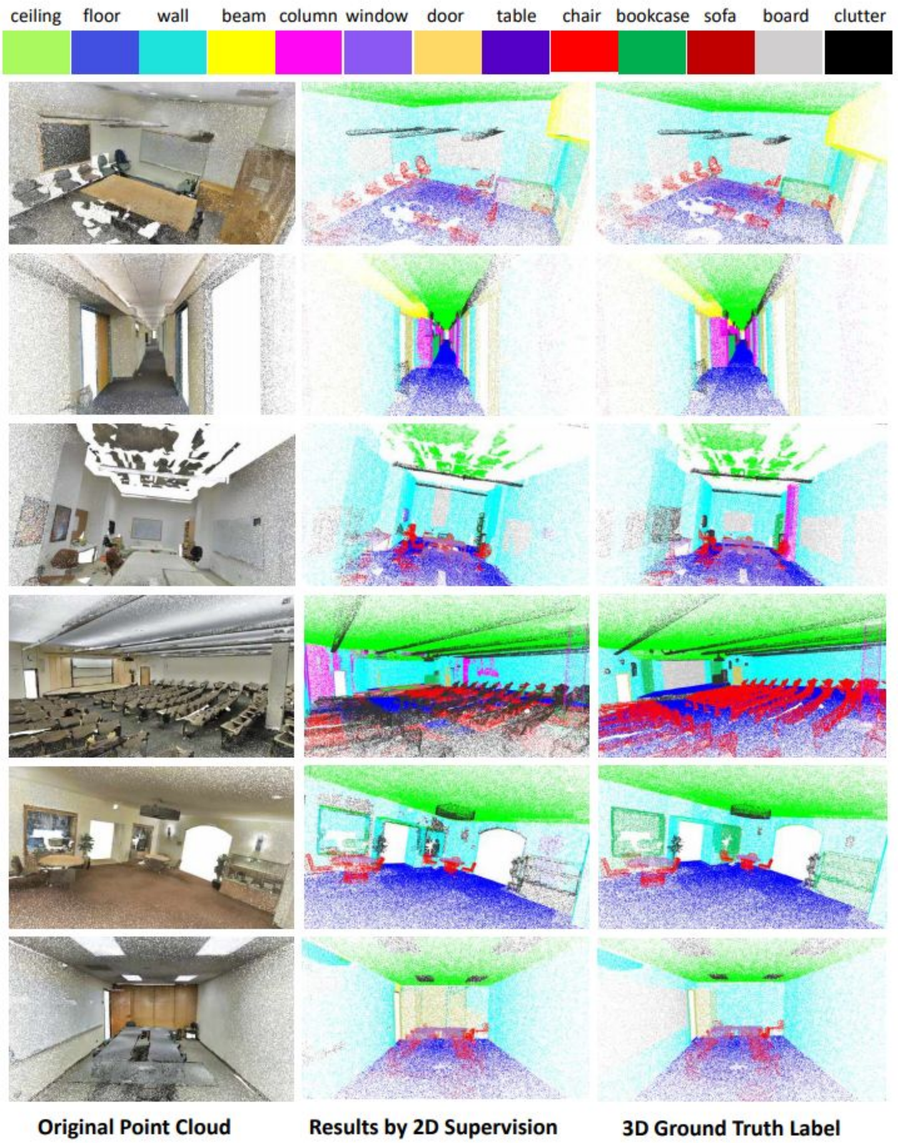
©2020 Arithmer.Inc. 当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。 Arithmer Confidential

Inference samples for SUNCG dataset



©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。
Arithmer Confidential

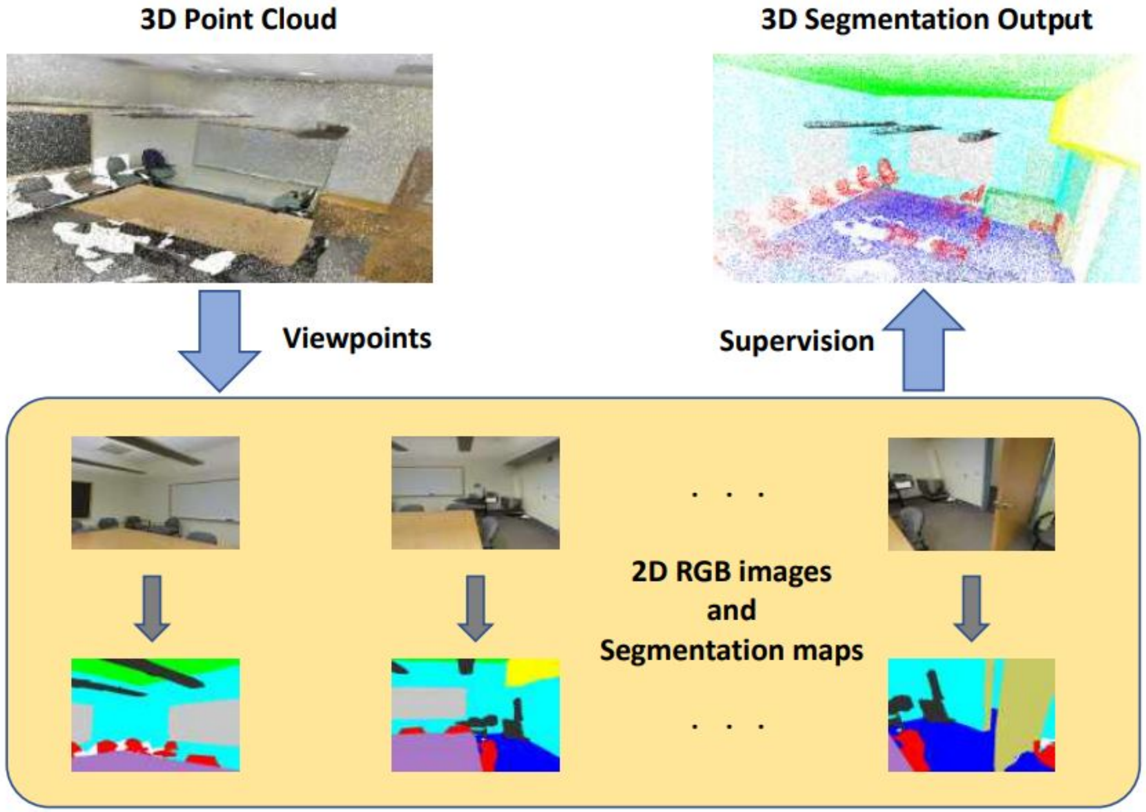
Inference samples for S3DIS dataset



©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。
Arithmer Confidential

Conclusion

- To the best of our knowledge, this is the **first work to apply 2D supervision for 3D semantic point cloud segmentation of wild scenes without using any 3D pointwise annotations.**
- Extensive experiments are conducted and the proposed method achieves comparable performance with the state-of-the-art 3D supervised methods on the popular SUNCG and S3DIS benchmarks.



©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。
Arithmer Confidential

Arithmer 株式会社
〒106-6040
東京都港区六本木一丁目６番１号 泉ガーデンタワー 38/40F(受付)
03-5579-6683
https://arithmer.co.jp/

Arithmer

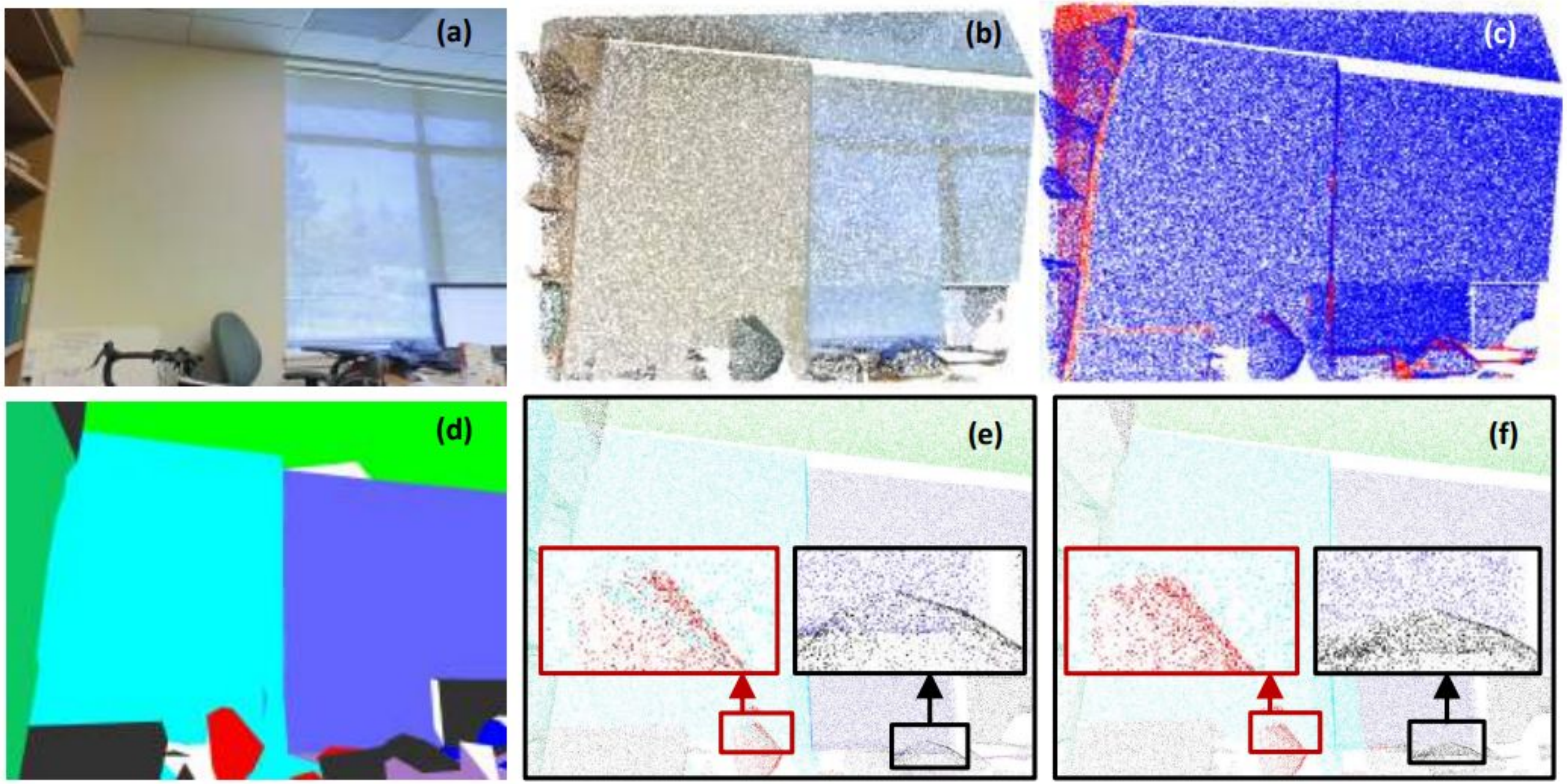
©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。
Arithmer Confidential

BACK UP

©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。
Arithmer Confidential

TITLE HERE

- Bullet list



©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。
Arithmer Confidential

Ablation study: projection method/OBSNet decoder

- Bullet list

Method	mAcc(%)	mIoU(%)	oAcc(%)
GPFN with DP (Ours)	61.9	45.0	73.4
GPFN with DP w/ D_v (Ours)	71.9	61.2	84.5
GPFN with PR w/o D_v (Ours)	65.3	50.8	79.1
GPFN with PR w/ D_v (Ours)	**87.3**	**70.37**	**91.8**

©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。
Arithmer Confidential

Ablation study - Encoder Design

- Bullet list

K-NN Graph	Pyramid	mAcc(%)	mIoU(%)	oAcc(%)
×	×	61.3	45.1	72.6
✓	×	65.1	48.6	78.4
×	✓	63.5	46.4	75.3
✓	✓	66.5	50.8	79.1

©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。
Arithmer Confidential

Ablation study - Amount of training data

- Bullet list

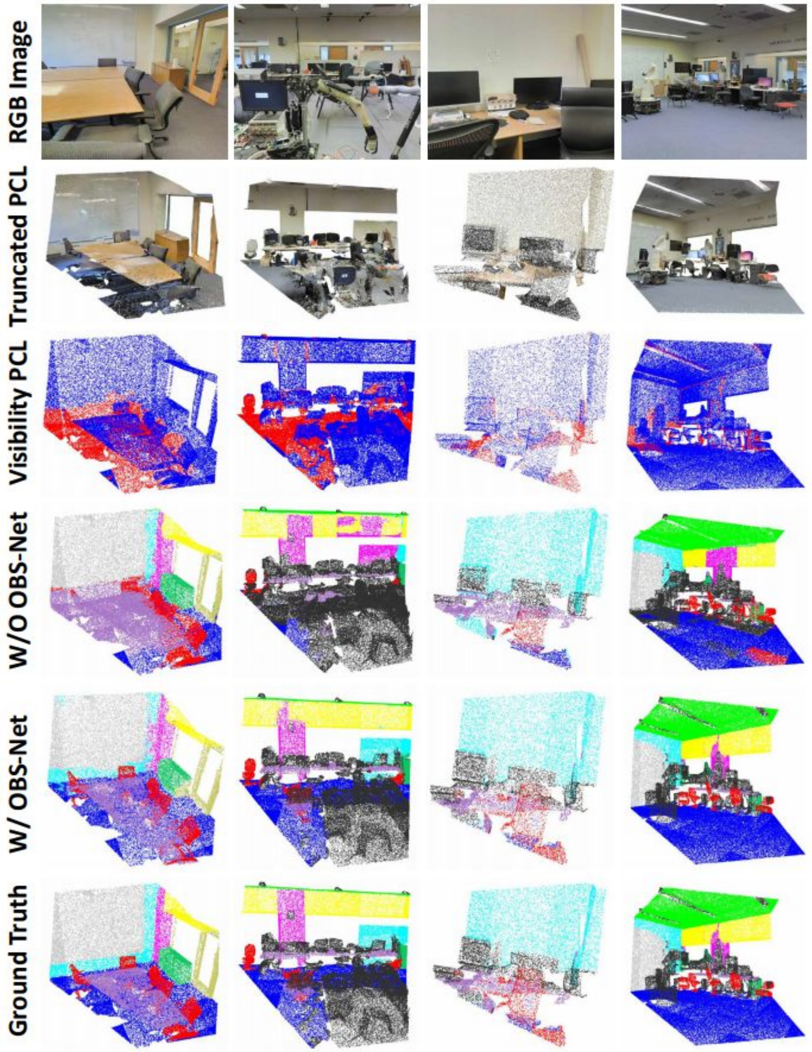
Training data	mAcc (%)	mIoU (%)	oAcc (%)
All	67.0	52.5	81.5
1/2	66.9	51.8	80.9
1/4	66.7	50.9	79.5
1/6	66.5	50.8	79.1
1/12	56.5	39.3	66.2
1/20	37.8	29.1	40.0

©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。 Arithmer Confidential

Ablation study - Visibility detection by OBSNet

- Bullet list

Dataset	Accuracy (%)					
	All	1/2	1/4	1/6	1/12	1/20
S3DIS	93.0	92.6	91.7	91.2	89.6	85.0



©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。　Arithmer Confidential

Transfer learning from synthesic to realistic dataset

- Bullet list

Training Data	mAcc(%)	mIoU(%)	oAcc(%)
Train Scratch on S3DIS	66.5%	50.8%	79.1%
Pretrained on SUNCG	67.0%	53.5%	81.3%

©2020 Arithmer.Inc.　当スライド（ページ）内の文章・画像等の無断転載はご遠慮ください。引用する場合には引用の要件を遵守してください。
Arithmer Confidential